

The University of Chicago

THE DEPENDENCE OF A FOCAL
POINT UPON CURVATURE IN THE
CALCULUS OF VARIATIONS

A DISSERTATION SUBMITTED TO THE
FACULTY OF THE DIVISION OF THE
PHYSICAL SCIENCES IN CANDIDACY FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF MATHEMATICS

1937

By

ALSTON SCOTT HOUSEHOLDER

Private Edition, Distributed by
THE UNIVERSITY OF CHICAGO LIBRARIES
CHICAGO, ILLINOIS

Reprinted from
CONTRIBUTIONS TO THE CALCULUS OF VARIATIONS, 1933-37
The University of Chicago Press, 1937

The University of Chicago

THE DEPENDENCE OF A FOCAL
POINT UPON CURVATURE IN THE
CALCULUS OF VARIATIONS

A DISSERTATION SUBMITTED TO THE
FACULTY OF THE DIVISION OF THE
PHYSICAL SCIENCES IN CANDIDACY FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF MATHEMATICS

1937

By

ALSTON SCOTT HOUSEHOLDER

Private Edition, Distributed by
THE UNIVERSITY OF CHICAGO LIBRARIES
CHICAGO, ILLINOIS

Reprinted from
CONTRIBUTIONS TO THE CALCULUS OF VARIATIONS, 1933-37
The University of Chicago Press, 1937



TABLE OF CONTENTS

Section	Page
Introduction	1
1. An invariant first curvature of a curve	2
2. The first and second variations of the integral	9
3. A second form of the Jacobi equations	19
4. The conjugate point	26
5. The focal point of a curve	32
6. The focal point of a variety	35



Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41548411_0106

THE DEPENDENCE OF A FOCAL POINT UPON CURVATURE
IN THE CALCULUS OF VARIATIONS

Introduction. This paper is concerned with problems of the calculus of variations in n -space in parametric form. It discusses the properties of the focal point of a curve D or of a variety V_N on an extremal arc E to which D or V_N is transversal. For the case of an arc D a tensor definition of curvature is described and it is shown that when the direction of D at its intersection P_0 with E is fixed the focal point P_2 of the arc D moves monotonically on E as the curvature of D at P_0 decreases from $+\infty$. It turns out that for certain directions of the arc D transversal to E at P_0 there is no focal point P_2 between P_0 and P_1 on the extremal arc E if the value of the curvature of D at P_0 is less than a properly chosen constant.

The methods used here are those of tensor analysis as developed by Cartan¹ for spaces of Finsler.²

In Section 1 an invariant first curvature of a curve D is defined and an interpretation is made of this curvature for a curve lying on a variety V_N . In Section 2 the first and second variations of the integral under consideration and certain tensors associated with them are discussed. In particular the Jacobi

¹E. Cartan, Les espaces de Finsler, Actualités scientifique and industrielles 79, Exposés de géométrie II (1934). Hereafter this monograph will be referred to simply as Cartan.

²P. Finsler, Über Kurven und Flächen in allgemeinen Räumen, Dissertation, Göttingen (1918).

equations of Cartan¹ analogous to those of Wren² are deduced. A system of orthogonal coordinates is developed in Section 3 whereby the orthogonal solutions of the Jacobi equations can be defined by a system of equations in $n - 1$ instead of n dependent variables. These equations are employed in Section 4 to derive certain properties of a point P_1 conjugate to P_0 of arbitrary order, and a matrix of functions $W_{ab}(s)$ is defined from their solutions, to be used in Section 5 to relate focal points to the invariant of curvature in the manner described in the first paragraph above. This is a generalization of the results obtained by Bliss³ for the plane and by White⁴ for 3-space. The point P_2 may approach P_1 as the curvature approaches $-\infty$, or it may attain the position of P_1 for some finite value of the curvature, depending upon the relation of the initial direction of the arc D at P_0 to the orthogonal solutions of the Jacobi equations which vanish at P_0 and at P_1 . Finally the focal point of a variety V_N is defined in Section 6 and some simple properties are deduced.

1. An invariant first curvature of a curve. Consider a parametric integral of a problem of the calculus of variations of the form

$$I = \int L(x^1, \dots, x^n, x^{1'}, \dots, x^{n'}) dt,$$

¹Cartan, p. 40.

²F. L. Wren, A new theory of parametric problems in the calculus of variations, Contributions to the calculus of variations 1930, University of Chicago, pp. 171 ff.

³G. A. Bliss, The second variation of a definite integral when one end-point is variable, Transactions of the American Mathematical Society, vol. 3 (1902), pp. 132-141.

⁴M. B. White, The dependence of focal points upon curvature for problems of the calculus of variations in space, Transactions of the American Mathematical Society, vol. 13 (1912), pp. 175 -198.

where the function $L(x, x')$ satisfies the conditions that for all elements (x, x') with x in a certain region R of x -space and $x' \neq 0$,

- a) L is analytic;
- b) L is positively homogeneous of the first degree in x' ;
- c) the quadratic form $g_{ij}(x, x')x^i x^j$, where

$$(1:1) \quad 2g_{ij}(x, x') = \partial^2 L^2(x, x') / \partial x^i \partial x^j,$$

is positive definite. From these assumptions it follows that $g_{ij}(x, x')$ is positively homogeneous of degree zero in x' , that $L^2(x, x') = g_{ij}x^i x^j$, and that L^2 and the determinant $g = |g_{ij}|$ are positive and non-null.¹ Moreover the functions g_{ij} are components of a covariant tensor if the x^i are components of a contravariant vector, as one may prove directly from the laws of transformation of these functions under an analytic coordinate transformation. This tensor defines the metric of a Finsler space in the following manner. If X^i and Y^j are components of arbitrary contravariant vectors defined at the element (x, x') formed by a point x and a direction x' , then the square of the length of X^i , and the angle θ between the two vectors, are defined by the equations

$$X^2 = g_{ij}(x, x')X^i X^j, \quad XY \cos \theta = g_{ij}(x, x')X^i Y^j.$$

Accordingly, if at every point of the curve D whose equations are

$$(1:2) \quad x^i = \phi^i(u),$$

there is defined a vector $x^i(u)$, whose components are not necessarily the derivatives of the functions $\phi^i(u)$, giving the direc-

¹Cartan, p. 10.

tion of the element (x, x') at that point, then the differential of the arc length of the curve at this element (x, x') is by definition given by the equation

$$ds^2 = g_{ij}(x, x') d\phi^i d\phi^j.$$

According to this definition, which is that of Finsler¹ and of Cartan², the arc D has an infinity of lengths, one for each choice of the one-parameter family of vectors $x^{i'}(u)$. Angles and lengths will be understood always in this sense as angles and lengths in Finsler space.

The fundamental contravariant tensor g^{ij} , whose determinant is g^{-1} , is defined just as in Riemann space by the equations $g_{ik}g^{kj} = \delta_i^j$. Likewise the process of raising and lowering indices on contraction with these two fundamental tensors is possible just as in Riemann space. Hence the contravariant components X^i and the associated covariant components $X_i = g_{ij}X^j$ may be regarded as being but different representations of the same vector and as defining the same direction in Finsler space. A like point of view is taken with regard to tensors; e.g., T_{ij} , $T_i^j = g^{jk}T_{ik}$, $T^i_j = g^{ik}T_{kj}$, $T^{ij} = g^{ik}T_k^j$ are to be regarded as different representations of the same tensor.

The contravariant vector $x^{i'}$ defining the direction of the element (x, x') may always be replaced by any positive multiple $kx^{i'}$ of itself in the tensor g_{ij} without affecting the metric, because of the homogeneity of the g_{ij} . If $k^{-1} = L(x, x')$, we obtain a vector

$$(1:3) \quad \rho^i = x^{i'}/L(x, x')$$

¹Finsler, loc. cit., pp. 32-33.

²Cartan, p. 4.

satisfying the equation

$$(1:4) \quad g_{ij} \ell^i \ell^j = L^2(x, \ell) = 1,$$

and which is, therefore, a unit vector. We observe that

$$(1:5) \quad \ell_i = g_{ij} \ell^j = \partial L(x, \ell) / \partial \ell^i,$$

a formula justifiable with the help of the fact that $2g_{ij} = \partial^2 L^2 / \partial x^i \partial x^j$ is positively homogeneous of order zero in the variables x^i . Hereafter ℓ will denote the unit vector in the direction of the element, as defined in equation (1:3).

The absolute differential of a tensor $T^{:::}(x, \ell)$ as defined by Cartan¹ is a tensor of the same type as the original; its components $DT^{:::}$ differ from the ordinary differentials $dT^{:::}$ by a linear form in the components $T^{:::}$. For a vector X ,

$$(1:6) \quad DX^i = dX^i + X^j \omega_j^i(d), \quad DX_i = dX_i - X_j \omega_i^j(d),$$

where the Pfaffians $\omega_i^j(d)$ are of the form

$$(1:7) \quad \omega_i^j(d) = \Gamma_i^j{}_k dx^k + C_i^j{}_k dx^{k'} = \Gamma_i^{*j}{}_k dx^k + A_i^j{}_k D\ell^k.$$

The coefficients Γ , C , Γ^* , and A are so determined by Cartan that the equations

$$(1:8) \quad Dg_{ij} = 0$$

hold identically. Their values are obtained by contracting the fundamental contravariant tensor g^{ij} with the functions defined by the equations²

¹Cartan, pp. 5, 17.

²Cartan, pp. 11, 13, 16.

$$\begin{aligned}
 2A_{ijk} &= 2LC_{ijk} = L \partial g_{ij} / \partial x^k, \\
 \Gamma_{ijk} &= (\partial g_{ij} / \partial x^k + \partial g_{jk} / \partial x^i - \partial g_{ki} / \partial x^j) / 2 \\
 (1:9) \quad &+ C_{ikm} \partial G^m / \partial x^j - C_{jkm} \partial G^m / \partial x^i, \\
 \Gamma_{ijk}^* &= \Gamma_{ijk} - C_{ijm} \partial G^m / \partial x^k,
 \end{aligned}$$

where

$$2G_m = x^{k'} \partial^2 f / \partial x^m \partial x^k - \partial f / \partial x^m, \quad f = L^2/2.$$

The functions A_{ijk} and Γ_{ijk}^* are homogeneous of degree zero; the first set is symmetric in all three indices, the second in the two lower ones. In view of these equations and the homogeneity of the g_{ij} it is evident that

$$(1:10) \quad \mathcal{L}^1 A_{ijk} = x^{1'} C_{ijk} = 0.$$

A vector X is said to be transported by parallelism along a curve C with respect to the directions $\mathcal{L}(t)$ of the elements $[x(t), \mathcal{L}(t)]$ taken along the curve when its absolute differential vanishes along C . Since the absolute differential operator obeys the same laws as the ordinary one, and is identical with this when operating upon a scalar invariant, it is a consequence of this definition of parallelism and of equations (1:8) that lengths and angles are invariant under parallel transport, just as in Riemann space. One needs only to observe that if X and Y are vectors undergoing parallel transport, then the differential of $g_{ij} X^i Y^j$ is zero.

Now consider an arc without singular points of a curve D whose equations are

$$(1:11) \quad x^i = \phi^i(\sigma),$$

with a unique element $[\phi(\sigma), \mathcal{L}(\sigma)]$ at every point whose di-

rection ℓ is normal to D:

$$(1:12) \quad \ell_1 \phi^{1'} = 0.$$

Let the parameter σ be chosen as the arc length of D with respect to the elements (ϕ, ℓ) , so that

$$(1:13) \quad g_{1j}(x, \ell) \phi^{1'} \phi^{j'} = 1.$$

Differentiating absolutely and using equations (1:8) we find the relation

$$(1:14) \quad g_{1j} \phi^{1'} D \phi^{j'} / d\sigma = 0,$$

which shows that the vector $D \phi^{j'} / d\sigma$ is normal to D. This is, by definition, the first normal to D, and its length $1/\rho$ is the first curvature. Let us use the notation n^1 for the unit vector in the direction of the first normal satisfying the equations

$$(1:15) \quad n^1 / \rho = D \phi^{1'} / d\sigma, \quad g_{1j} n^1 n^j = 1.$$

If $R \ell^1$ is the vector in the direction of the element whose projection upon the first normal is of magnitude ρ , the reciprocal $1/R$ of its length is a scalar invariant which may be called the principal curvature of D at the normal element (ϕ, ℓ) :

$$(1:16) \quad 1/R = \ell_1 n^1 / \rho.$$

This invariant has the following important properties:

THEOREM 1:1. If u is an arbitrary parameter along a curve D, the principal curvature $1/R$ of D is given by the equation

$$(1:17) \quad 1/R = \ell_1 D^2 x^1 / du^2 / g_{1j} (dx^1 / du) (dx^j / du),$$

and its value does not involve the derivative $D \ell^1 / du$. In the equation (1:17) the symbol $D^2 x^1 / du^2$ stands for $D(dx^1 / du) / du$.

To prove this observe that

$$n^1/\rho = D^2x^1/d\sigma^2 = (D^2x^1/du^2)(du/d\sigma)^2 + (dx^1/du)(d^2u/d\sigma^2).$$

Then from equations (1:16) and (1:12) it follows that

$$1/R = \mathcal{L}_1(D^2x^1/du^2)(du/d\sigma)^2.$$

Hence equation (1:17) follows from the definition of the differential of arc length. To complete the proof of the theorem, we find from equations (1:6), (1:7) and (1:10) that the term involving $D\mathcal{L}^1/du$ in the expression for D^2x^1/du^2 vanishes on contracting with $\mathcal{L}_1$.

In case the curve D lies on a variety V_N defined by equations of the form

$$(1:18) \quad x^1 = x^1(y^1, \dots, y^N),$$

the curve can be defined by equations of the form

$$(1:19) \quad y^p = y^p(u) \quad (p = 1, \dots, N).$$

Let the direction $\mathcal{L}$ of the element $(x, \mathcal{L})$ be defined and normal to V_N at every point of a region on V_N . Introduce the notations

$$(1:20) \quad \begin{aligned} x^1_p &= \partial x^1 / \partial y^p, \\ x^1_{pq} &= D x^1_p / \partial y^q \\ &= \partial^2 x^1 / \partial y^q \partial y^p + x^j_p (\Gamma^*_{j k} x^k_q + A^1_{j k} D \mathcal{L}^k / \partial y^q). \end{aligned}$$

Then the coefficients of the first and second fundamental forms belonging to V_N are defined as in Riemann space by the formulas

$$(1:21) \quad g_{pq} = g_{ij} x^i_p x^j_q, \quad L_{pq} = \mathcal{L}_1 x^1_{pq}.$$

Because of the symmetry of the g_{ij} and the $\Gamma^*_{i j}{}^k$ in i and j , we may apply equations (1:10) to obtain

LEMMA 1:1. The coefficients, g_{pq} and L_{pq} , of the first and second fundamental forms at a normal element (x, ℓ) of a variety V_N in Finsler space are symmetric in their indices and depend only upon the direction of the element at that point when the point itself is fixed.

Using equation (1:17), the principal curvature of the curve D can be calculated. We have

$$D^2 x^1 / du^2 = x^1_{py} y^p y'' + x^1_{pq} y^p y'^q,$$

so that

$$\ell_1 D^2 x^1 / du = L_{pq} y^p y'^q,$$

since the element is normal to the variety. Hence

$$(1:22) \quad 1/R = L_{pq} dy^p dy^q / g_{pq} dy^p dy^q.$$

Then on applying the lemma we deduce easily

THEOREM 1:2. The principal curvature of a curve D on a variety V_N in Finsler space at an element (x, ℓ) for which the direction ℓ is normal to V_N , is given by (1:22). This expression does not involve the differential $D\ell^1$ and hence is independent of the directions ℓ of neighboring elements on V_N . It has the same value, when the vector ℓ is given, for all curves on V_N through the point x and having the same tangent at x .

2. The first and second variations of the integral. Consider a one-parameter family of curves C with equations of the form

$$(2:1) \quad x^i = x^i(t, u) \quad [t_1(u) \leq t \leq t_2(u)],$$

where the functions $x^i(t, u)$, $t_1(u)$ and $t_2(u)$ are of class C^n . We suppose the curves C to lie entirely within the region R of

x -space where the properties of the function $L(x, x')$ are defined, and to have their endpoints varying on two curves D_1 and D_2 with equations

$$(2:2) \quad x^1 = x^1[t_1(u), u],$$

$$(2:3) \quad x^1 = x^1[t_2(u), u],$$

respectively. Either of these curves D may reduce to a point. Define the vector ℓ of the element (x, ℓ) and the variation ξ by the equations

$$(2:4) \quad x^{1'} = \partial x^1 / \partial t = L(x, x') \ell^1, \quad \xi^1(t, u) = \partial x^1 / \partial u.$$

Then ξ^1 is a vector. For the partial absolute derivatives with respect to t and u and the element (x, ℓ) the equations

$$(2:5) \quad D \xi^1 / \partial t = \partial^2 x^1 / \partial t \partial u + \xi^j (\Gamma_j^1 k^{x^k} + C_j^1 k^{x^k}),$$

and, after applying (1:10),

$$(2:6) \quad D x^{1'} / \partial u = \partial^2 x^1 / \partial u \partial t + x^{k'} \Gamma_k^1 \xi^j,$$

hold along each curve C . From these we find

$$(2:7) \quad D x^{1'} / \partial u = D \xi^1 / \partial t - L \xi^j C_j^1 k^{x^k} D \ell^k / \partial t$$

with the help of (1:10) and (1:6) above and equation (13) in Cartan.

Now

$$L^2 = g_{1j} x^{1'} x^{j'},$$

$$2L \partial L / \partial u = \partial(g_{1j} x^{1'} x^{j'}) / \partial u = 2g_{1j} x^{1'} D x^{j'} / \partial u$$

since the ordinary and absolute derivatives are identical for a scalar. From this result and (2:7), (2:5) and (2:6), and by applying (1:10),

$$(2:8) \quad L \partial L / \partial u = g_{1j} x^{1j} D \xi^j / \partial t,$$

or, by (2:4),

$$(2:9) \quad \partial L / \partial u = \ell_1 D \xi^1 / \partial t.$$

From these results we now deduce the following theorem:

THEOREM 2:1. The integral

$$I(u) = \int_{t_1(u)}^{t_2(u)} L[x(t, u), x'(t, u)] dt,$$

taken along the family $x^1 = x^1(t, u)$ $[t_1(u) \leq t \leq t_2(u)]$, has as its first derivative the value

$$I'(u) = [L dt/du]_{t_1}^{t_2} + \int_{t_1}^{t_2} \ell_1 (D \xi^1 / \partial t) dt.$$

Since

$$\partial(g_{1j} \ell^1 \xi^j) / \partial t = g_{1j} \xi^1 D \ell^j / \partial t + \ell_1 D \xi^1 / \partial t,$$

we may transform the integral in the first derivative $I'(u)$ by parts and obtain

$$(2:10) \quad I'(u) = [L dt/du + \ell_1 \xi^1]_{t_1}^{t_2} - \int_{t_1}^{t_2} (D \ell_1 / \partial t) \xi^1 dt.$$

From this equation one finds in the usual way, with the help of the fundamental lemma of the calculus of variations, the following theorem:

THEOREM 2:2. For the problem of minimizing the integral I in a class of arcs joining two curves D_1 and D_2 a minimizing arc of class C'' must be a solution of the equations

$$D \ell^1 / dt = 0,$$

where the vector ℓ^1 has the direction tangential to C . The

solutions of this equation are the extremals of the problem.¹

Now suppose that the family of curves C contains for $u = 0$ an extremal C_0 which minimizes the integral I with respect to neighboring curves of class C'' joining D_1 and D_2 , and hence which causes the first variation $I'(0)$ to vanish for all families of the form (2:1) joining D_1 and D_2 . Then from equation (2:10) the first variation $I'(0)$ along the extremal arc C_0 has the value

$$(2:11) \quad I'(0) = [Ldt/du + \mathcal{L}_1 \xi^1]_{t_1}^{t_2}.$$

If we let $t(u)$ stand for either $t_1(u)$ or $t_2(u)$, then

$$(2:12) \quad dx^1/du = x^1 dt/du + \xi^1.$$

If, furthermore, the family (2:1) is specialized to join the intersection point of D_1 and C_0 to the arc D_2 , and if the variable t is the value of the integral I taken from the intersection of C_0 and D_1 along each curve C , then $t_1(u) = 0$ identically, $\xi^1(0, u) = 0$, and $t_2(u) = I(u)$. Hence if C_0 is to furnish a minimum for the integral in the class of arcs joining the fixed curves D_1 and D_2 , then $t_2'(0) = I'(0) = 0$. Thus at the intersection point P_2 of C_0 with D_2 , we find from (2:11) and (2:12), since $dt/du = 0$, that

$$\mathcal{L}_1 \xi^1 = \mathcal{L}_1 dx^1/du = 0,$$

and C_0 and D_2 are therefore orthogonal with respect to the element $(x, \mathcal{L})$ tangent to C_0 at P_2 . In an analogous manner we find that C_0 must be orthogonal to D_1 at their intersection point P_1 .

A similar argument shows that when the arc C_0 is orthogonal to D_1 it is also orthogonal to every curve $t = \text{constant}$ when

¹The equations of the extremals were deduced differently in Cartan, p. 17.

the parameter t is chosen along each arc C to be the value of the integral I measured from D_1 . The equation $\mathcal{L}_1 \xi^1 = 0$ is thus an identity in t along C_0 , when the choice of parameter t has been so made, and the absolute derivative of the expression $\mathcal{L}_1 \xi^1$ must also vanish identically in t along C_0 as indicated by the equations

$$(2:13) \quad \mathcal{L}_1 \xi^1 = \mathcal{L}_1 D \xi^1 / \partial t = 0.$$

THEOREM 2:3. If an arc C_0 is an extremal minimizing the integral I with respect to the arcs C of class C'' joining two curves D_1 and D_2 of class C'' , then D_1 and D_2 are orthogonal to C_0 with respect to the elements $(x, \mathcal{L})$ along C_0 whose directions are tangent to C_0 at its intersections with D_1 and D_2 .

THEOREM 2:4. Let a one-parameter family of curves C of class C'' contain an extremal C_0 cut orthogonally by the curve D_1 . Then any curve D , to which the value of the integral from D_1 along the curves C is constant, is also orthogonal to C_0 with respect to the elements $(x, \mathcal{L})$ which have directions tangent to the curves C .

Returning to an arbitrary family of arcs C joining the curves D_1 and D_2 with an arbitrarily chosen parameter t , let us calculate the second variation of the integral $I(u)$. Since

$$L \partial^2 L / \partial u^2 = \partial(L \partial L / \partial u) / \partial u - (\partial L / \partial u)^2,$$

we have from (2:8), (2:9), (1:10) and the fact that the absolute and ordinary derivatives of a scalar are equal,

$$L \partial^2 L / \partial u^2 = g_{1j} (D x^{1j} / \partial u) (D \xi^j / \partial t) + g_{1j} x^{1j} D (D \xi^j / \partial t) / \partial u - (\mathcal{L}_1 D \xi^1 / \partial t)^2.$$

Hence

$$\begin{aligned}
 L \partial^2 L / \partial u^2 &= (g_{1j} - \ell_1 \ell_j)(D\xi^1 / \partial t)(D\xi^j / \partial t) \\
 (2:14) \quad &- L g_{kj1} \xi^k (D\xi^j / \partial t)(D\ell^1 / \partial t) \\
 &+ g_{1j} x^{1'} D(D\xi^j / \partial t) / \partial u,
 \end{aligned}$$

as one sees from (2:7).

Let the symbols d and D signify the differentials, ordinary and absolute, respectively, with respect to the parameter t of our family, while δ and Δ signify the differentials with respect to the parameter u . Then consider the last term on the right in (2:14). First we have, from equations (1:6) and (1:8)

$$\begin{aligned}
 \Delta D \xi^j &= \delta d \xi^j + d \xi^h \omega_h^j(\delta) + \delta \xi^h \omega_h^j(d) \\
 &+ \xi^k [\delta \omega_k^j(d) + \omega_k^{h(d)} \omega_h^j(\delta)].
 \end{aligned}$$

Interchanging the order of differentiation and subtracting we have

$$\begin{aligned}
 (\Delta D - D \Delta) \xi^j &= \xi^k \{ \delta \omega_k^j(d) - d \omega_k^j(\delta) - \omega_k^{h(\delta)} \omega_h^j(d) \\
 &+ \omega_k^{h(d)} \omega_h^j(\delta) \} \\
 &= \xi^k \{ (\omega_k^j)' - [\omega_k^{h(d)} \omega_h^j] \},
 \end{aligned}$$

where

$$\begin{aligned}
 (\omega_k^j)' &= \delta \omega_k^j(d) - d \omega_k^j(\delta) \\
 [\omega_k^{h(d)} \omega_h^j] &= \omega_k^{h(\delta)} \omega_h^j(d) - \omega_k^{h(d)} \omega_h^j(\delta)
 \end{aligned}$$

in the symbolism of Cartan.¹ The quantity within the round brackets is the Pfaffian which Cartan² denotes by $\Omega_k^j(d, \delta)$, and with this notation we have

¹E. Cartan, Leçons sur la géométrie des espaces de Riemann (1928), pp. 204 ff.

²Cartan, pp. 32 ff.

$$(2:15) \quad (\Delta D - D\Delta) \xi^j = \xi^k \Omega_k^j(d, \delta).$$

For an integration by parts we need the formula

$$d(\ell_1 \Delta \xi^i) = \ell_1 D \Delta \xi^i + g_{1j} D \ell^i \Delta \xi^j.$$

After such an integration the general expression for the second variation of the integral $I(u)$ along an arbitrary curve C is found from the result of Theorem 2:1, and with the help of (2:9), (2:14), (2:15), and the property $\Omega_{1j} = -\Omega_{j1}$, to have the value

$$(2:16) \quad \begin{aligned} I''(u) = & [d(Ldt/du)/du + \ell_1(D\xi^1/\partial t)(dt/du) + \ell_1 D\xi^1/\partial u]_{t_1}^{t_2} \\ & + \int_{t_1}^{t_2} \{ L^{-1}(g_{1j} - \ell_1 \ell_j)(D\xi^1/\partial t)(D\xi^j/\partial t) \\ & - \ell^1 \xi^j \Omega_{1j}(d, \delta)/\partial t \partial u - (g_{1j} D\xi^j/\partial u \\ & + c_{1jk} \xi^j D\xi^k/\partial t) D\ell^1/\partial t \} dt. \end{aligned}$$

Now suppose that C_0 is an extremal and that the parameter t is the value of the integral I taken along the arcs C from D_1 . Then for $u = 0$ in (2:16) the last group of terms under the integral sign drop out because of Theorem 2:2, and the terms in $\ell_1 D\xi^1/\partial t$ vanish because of (2:13). From equation (41) in Cartan contracted with the vectors $\ell^i \xi^j$ we find

$$\begin{aligned} \ell^1 \xi^j \Omega_{1j}(d, \delta)/\partial t \partial u &= \ell^1 \xi^j \{ s_{1jkh} (D\ell^k/\partial t)(D\ell^h/\partial u) \\ &+ p_{1jkh} (\ell^k D\ell^h/\partial u - \xi^k D\ell^h/\partial t) + r_{1jkh} \ell^k \xi^h \} \\ &= \xi^j (p_{ojoh} D\ell^h/\partial u + r_{ojoh} \xi^h), \end{aligned}$$

where the subscripts 0 signify contraction with the vector ℓ .

In securing this result use is made of the fact that $D\ell^1/\partial t = 0$ along the extremal C_0 , and of the formula $\ell^1 s_{1jkh} = 0$ which is a

consequence of (1:10) above and equation (XVI) in Cartan. This formula (XVI) is the definition of the tensor S in terms of the tensor A in (1:9) above. The tensor P is defined on p. 35 of Cartan in terms of the tensor C and the expressions Γ^* of (1:9). The tensor R is defined in terms of these same quantities and forms a generalization of the Riemann tensor, as given on p. 36 of Cartan. But it follows from (1:10) above, equation (XVII) of Cartan, and the definition of the tensor A_{ijk} in Cartan Section VII, that $P_{0j0h} = 0$. Hence

$$\ell^i \xi^j \Omega_{ij}(d, \delta) / \partial t \partial u = \xi^i \xi^j R_{0i0j},$$

and the second variation along C_0 becomes

$$\begin{aligned} I''(0) &= [d^2t/du^2]^{t_2} + [\ell_1 D \xi^1 / \partial u]_{t_1}^{t_2} \\ (2:17) \quad &+ \int_{t_1}^{t_2} \{ g_{1j}(D \xi^1 / \partial t)(D \xi^j / \partial t) - R_{0i0j} \xi^i \xi^j \} dt, \end{aligned}$$

since $dt/du = 0$ at D_1 and at the intersection of C_0 with D_2 , and on account of our choice of parameter giving $L = 1$ on the arcs C .

We wish a further transformation of the part outside the integral sign, however. By differentiating the equations (2:12) we find

$$\begin{aligned} D(dx^i/du)/du &= (Dx^i / \partial t)(dt/du)^2 + (Dx^i / \partial u + D \xi^i / \partial t)dt/du \\ &+ x^i d^2t/du^2 + D \xi^i / \partial u. \end{aligned}$$

But if $t_1(u) = 0$ along D_1 , and $t_2(u) = I(u)$ along D_2 , then from the last formula

$$\begin{aligned} D(dx^i/du)/du|_{t=u=0} &= D \xi^i / \partial u|_{t=u=0} \\ D(dx^i/du)/du|_{t=t_2}^{u=0} &= [\ell^i I''(0) + D \xi^i / \partial u]_{t=t_2}^{u=0}, \\ \ell_1 D(dx^i/du)/du|_{t=t_2}^{u=0} &= I''(0) + [\ell_1 D \xi^i / \partial u]_{t=t_2}^{u=0}. \end{aligned}$$

The second member of the last equation is the sum of the first two terms in the second member of (2:17) since $t_2''(0) = I''(0)$. Now using the expression (1:17) for the principal curvature of the arc D_1 or D_2 , and using the fact that $dx^i/du = \xi^i$ in (2:12) since $dt/du = D$ at either end-point of C_0 we find that we may write in place of (2:17), the equation

$$(2:18) \quad I''(0) = R^{-1} g_{ij} \xi^i \xi^j \Big|_{t_1}^{t_2} + \int_{t_1}^{t_2} \{ g_{ij} (D\xi^i/\partial t) (D\xi^j/\partial t) - R_{oioj} \xi^i \xi^j \} dt.$$

THEOREM 2:5. If C_0 is an extremal for the integral I , cut transversally by the curves D_1 and D_2 , if the parameter t is the generalized arc length of C_0 with respect to elements (x, ℓ) whose directions are tangential to C_0 as defined in Section 1 above, and if furthermore the variations $\xi^i(t)$ are orthogonal variations along C_0 satisfying the condition $\rho_i \xi^i = 0$, then the second variation of the integral I taken along C_0 between the curves D_1 and D_2 is given by (2:17) or (2:18). The factor R^{-1} takes the values of the principal curvatures of D_1 and D_2 at t_1 and t_2 , respectively, and R_{oioj} is defined in terms of the generalized Riemann tensor R_{hikj} by the formula $R_{oioj} = \rho^h \rho^k R_{hikj}$

Finally, let us calculate the Euler equations for the integral of the second variation. These are the Jacobi equations of the original problem. Replace the parameter t by s to indicate that the generalized arc length is being used and denote, as usual, the integrand in (2:18) by

$$(2:19) \quad 2\Omega(x, \xi, \xi') = g_{ij} (D\xi^i/ds) (D\xi^j/ds) - R_{oioj} \xi^i \xi^j.$$

It is permissible to use total derivatives with respect to s since we are now concerned with the single extremal C_0 only.

Along this extremal, by (1:6), (1:7) and Theorem 2:2, we have

$$D\xi^i/ds = \xi^{i'} + \xi^k \Gamma_{kh}^i \rho^h.$$

Then because of the symmetry¹ of g_{ij} and R_{oioj} in i and j , the partial derivatives of Ω with respect to ξ^i and $\xi^{i'}$ have the values

$$\begin{aligned} \Omega_{\xi^i} &= g_{ij} D\xi^j/ds, \\ (2:20) \quad \Omega_{\xi^{i'}} &= \rho^j \Gamma_{ikj}^* D\xi^k/ds - R_{oioj} \xi^j. \end{aligned}$$

Since Ω_{ξ^i} appears as a covariant vector we may again employ successively (1:6), (1:7), Theorem 2:2, and the first equation (2:20) with (1:8) to secure the equations

$$\begin{aligned} d\Omega_{\xi^i}/ds &= D\Omega_{\xi^i}/ds + \Omega_{\xi^j} \Gamma_{ik}^j \rho^k \\ &= g_{ij} D^2\xi^j/ds^2 + \Gamma_{ijk}^* \rho^k D\xi^j/ds. \end{aligned}$$

From this result and the second equation (2:20)

$$(2:21) \quad d\Omega_{\xi^i}/ds - \Omega_{\xi^{i'}} = g_{ij} D^2\xi^j/ds^2 + R_{oioj} \xi^j.$$

THEOREM 2:6. For orthogonal variations ξ^i satisfying the condition $\rho_i \xi^i = 0$ along an extremal arc C_0 the Euler equations for the integral of the second variation (2:13) along C_0 , which are the Jacobi equations along C_0 for the original integral I , are the equations

$$(2:22) \quad g_{ij} D^2\xi^j/ds^2 + R_{oioj} \xi^j = 0,$$

or the equivalent system

¹Cartan, (XXIII).

$$(2:23) \quad D^2 \xi^1 / ds^2 + R_{0 \ 0j}^1 \xi^j = 0.$$

This is the form given by Cartan.¹

If ξ^1 is an orthogonal variation satisfying the Jacobi equations (2:23) and defining directions at the endpoints of C_0 tangent to the curves D_1 and D_s , then with the help of (2:23) and the definition (2:19) of 2Ω we find that

$$\begin{aligned} d(g_{1j} \xi^1 D \xi^j / ds) / ds \\ = g_{1j} (D \xi^1 / ds) (D \xi^j / ds) + g_{1j} \xi^1 D^2 \xi^j / ds^2 = 2\Omega. \end{aligned}$$

In this case the integration in (2:18) can be effected and we find the following result:

THEOREM 2:7. For a set of orthogonal variations ξ^1 satisfying the Jacobi equations (2:23) along an extremal arc C_0 the second variation has the value

$$(2:24) \quad I''(0) = [R^{-1} g_{1j} \xi^1 \xi^j + g_{1j} \xi^1 D \xi^j / ds]_{s_1}^{s_2}.$$

3. A second form of the Jacobi equations. Consider an extremal arc E for the calculus of variations integral I , with equations of the form

$$(3:1) \quad x^1 = x^1(s) \quad (0 \leq s \leq S),$$

and let the generalized arc length be taken for the parameter s . Let the element (x, ρ) of Finsler space along this arc have the

¹Cartan, p. 40. The first member of equation (2:22) is the expression $K_1(\eta)$ of Wren with the notation η^1 replaced by ξ^1 . See his Dissertation to which reference has been made above. According to Wren the usual Jacobi expressions $J_1(\xi)$ are related to the $K_1(\xi)$ by the formula $J_1(\xi) = K_1(\xi) - (\rho_j \xi^j)'' \rho_1$, and furthermore $0 = \rho^1 J_1(\xi) = \rho^1 K_1(\xi) - (\rho_1 \xi^1)''$. Consequently a solution ξ^1 of $K_1(\xi) = 0$ makes $(\rho_1 \xi^1)'' = 0$ and will be an orthogonal solution of $J_1(\xi) = 0$ or $K_1(\xi) = 0$ if $\rho_1 \xi^1$ and $(\rho_1 \xi^1)'$ vanish at a single parameter value s_0 .

direction $x^{i'}(s)$. Then $L^2 = g_{ij}x^{i'}x^{j'} = 1$, so that from the definition (1:3) of the vector $\mathcal{L}^i$ it follows that

$$(3:2) \quad \mathcal{L}^i(s) = x^{i'}(s)$$

along E . Let P_0 denote the initial point $x^i(0)$ of the extremal arc.

At P_0 choose $n - 1$ unit vectors Y_a^i ($a = 1, \dots, n - 1$), orthogonal to E and to each other, and hence satisfying the equations

$$(3:3) \quad g_{ij}Y_a^iY_b^j = \delta_{ab}, \quad \mathcal{L}_iY_a^i = 0,$$

where $\delta_{ab} = \delta_a^b$ is the Kronecker delta. If these vectors are transported by parallelism along E , i.e. if the vectors $Y_a^i(s)$ are chosen all along E to satisfy

$$(3:4) \quad DY_a^i(s)/ds = 0,$$

then equations (3:3) are identities in s since lengths and angles are invariant under parallel transport.

Any vector $\xi^i(s)$ orthogonal to E at every point of E is expressible as a linear combination of the vectors Y_a^i , with coefficients which are scalar functions of s , in the form

$$(3:5) \quad \xi^i(s) = h_a(s)Y_a^i(s), \quad D\xi^i(s)/ds = h_a'(s)Y_a^i(s).$$

Because of the fact that absolute differentiation obeys the same laws as the ordinary differentiation and is identical with it when operating upon a scalar, the second set of equations in (3:5) is a consequence of the first and of (3:4). Contraction in (3:5) with $g_{ij}Y_b^j$ and application of (3:3) gives, after a change of indices,

$$(3:6) \quad h_a = g_{ij} y_a^i \xi^j, \quad h_a' = g_{ij} y_a^i D \xi^j / ds.$$

If the ξ^i satisfy the Jacobi equations (2:23) for the extremal arc E, then on differentiating the second of equations (3:6) and applying (1:8), (3:4), (2:23), and (3:5) we find

$$h_a'' = -R_{oioj} y_a^i y_b^j h_b.$$

In view of the symmetry of R_{oioj} in i and j referred to in the derivation of equations (2:20) we may set

$$(3:7) \quad P_{ab} = P_{ba} = R_{oioj} y_a^i y_b^j,$$

and have

$$(3:8) \quad h_a'' + P_{ab} h_b = 0$$

as a necessary condition for the vector ξ^i defined in (3:5) to satisfy the Jacobi equations.

This condition is also sufficient, as one may see with the help of the following lemma:

LEMMA 3:1. If the vectors y_a^i are chosen to satisfy equations (3:3) identically in s , then

$$y_a^i y_a^j = g^{ij} - \ell^i \ell^j.$$

To prove this assertion we set

$$y_{a1} = g_{ij} y_a^j.$$

Then equations (3:3) can be written

$$y_{a1} y_b^1 = \delta_{ab}, \quad \ell_1 y_a^1 = 0,$$

while these equations, together with

$$y_{a1} \ell^1 = 0, \quad \ell_1 \ell^1 = 1$$

show that the matrices

$$\begin{pmatrix} Y_{a1} \\ \ell_1 \end{pmatrix} \quad \text{and} \quad (Y_a^1 \quad \ell^1)$$

are reciprocal. Hence

$$Y_{a1} Y_a^j + \ell_1 \ell^j = \delta_1^j,$$

or

$$g_{1k} (Y_a^k Y_a^j + \ell^k \ell^j) = \delta_1^j.$$

On contracting both sides with g^{h1} we find

$$Y_a^h Y_a^j + \ell^h \ell^j = g^{hj}$$

which is equivalent to the desired identity.

Now to prove that the condition (3:8) is sufficient for the vector $\xi^1(s)$ in the first of equations (3:5) to be a solution of the Jacobi equations, we first find from equations (3:5) and (3:8) that

$$\begin{aligned} D^2 \xi^1 / ds^2 &= -P_{ab} h_b Y_a^1 \\ &= -R_{ojok} Y_a^j Y_b^k Y_a^1 h_b \\ &= -\xi^k R_{ojck} Y_a^j Y_a^1. \end{aligned}$$

Application of Lemma 3:1 yields

$$\begin{aligned} D^2 \xi^1 / ds^2 &= -\xi^h R_{ojoh} (g^{j1} - \ell^j \ell^1) \\ &= -\xi^h (R_o^1{}_{oh} - R_{oooh} \ell^1). \end{aligned}$$

However R_{1jkh} is skew symmetric in the indices k and h so that $R_{1joo} = 0$ and hence $R_{ojoo} = 0$. Hence if we contract the fourth of identities (XXIII) in Cartan with ℓ^j we find that $R_{oooh} = 0$, and this proves the sufficiency.

THEOREM 3:1. Let $Y_a^1(s)$ be a set of $n - 1$ unit vectors defined along an extremal arc E , each one auto-parallel along E and orthogonal to E and to the other vectors. Then every vector $\xi^1(s)$ orthogonal to E is expressible as a linear combination (3:5) of the vectors Y_a^1 with coefficients which are scalar invariant functions $h_a(s)$ of s . Moreover the vector $\xi^1(s)$ is a solution of the Jacobi equations for E if and only if these coefficients h_a satisfy the set of second order linear differential equations (3:8).

Let h_a and k_a be two sets of solutions of the equations (3:8), so that

$$h_a'' = -P_{ab}h_b, \quad k_a'' = -P_{ab}k_b.$$

Then in view of the symmetry of the P_{ab} ,

$$d(h_a k_a' - h_a' k_a)/ds = -h_a P_{ab} k_b + k_a P_{ab} h_b = 0.$$

Hence

LEMMA 3:2. If h_a and k_a are two sets of solutions of (3:8), then

$$(3:9) \quad h_a k_a' - h_a' k_a = \text{const.}$$

Let us then choose a fundamental group of $2(n - 1)$ sets of solutions h_{ab} and k_{ab} , of (3:8), to satisfy the initial conditions

$$(3:10) \quad h_{ab}(0) = k_{ab}'(0) = \delta_{ab}, \quad h_{ab}'(0) = k_{ab}(0) = 0.$$

It is understood that for a fixed b , h_{ab} and k_{ab} are two sets of solutions. Applying the lemma to these solutions we obtain the following sets of identities:

$$\begin{aligned}
 (3:11) \quad & h_{ca}k_{cb}' - h_{ca}'k_{cb} = \delta_{ab}, \\
 & h_{ca}h_{cb}' - h_{ca}'h_{cb} = 0, \\
 & k_{ca}k_{cb}' - k_{ca}'k_{cb} = 0;
 \end{aligned}$$

and these are equivalent to the set

$$\begin{aligned}
 (3:12) \quad & h_{ac}k_{bc}' - k_{ac}h_{bc}' = \delta_{ab}, \\
 & h_{ac}'k_{bc}' - k_{ac}'h_{bc}' = 0, \\
 & h_{ac}k_{bc} - k_{ac}h_{bc} = 0.
 \end{aligned}$$

To prove this equivalence let h, h', k, k' represent the matrices $(h_{ab}), (h_{ab}'), (k_{ab}), (k_{ab}')$; and let $\bar{h}$, etc., denote the transposes of these same matrices. Then equations (3:11) are equivalent to the single matrix equation

$$\begin{pmatrix} \bar{k}' & -\bar{k} \\ -\bar{h}' & \bar{h} \end{pmatrix} \begin{pmatrix} h & k \\ h' & k' \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}.$$

But this matrix equation remains valid if the factors in the left member are commuted, and if this is done equations (3:12) result.

LEMMA 3:3. If h_{ab} and k_{ab} are $2(n-1)$ sets of solutions of the differential equations (3:8), chosen to satisfy at any point the conditions (3:10), then these solutions satisfy the equivalent sets of identities (3:11) and (3:12).

Now consider the determinant

$$(3:13) \quad K = |k_{ab}|,$$

and let K_{ab} be the quotient of the cofactor of k_{ba} in K divided by K itself, as is required by the equations

$$(3:14) \quad K_{ac}k_{cb} = k_{ac}K_{cb} = \delta_{ab}.$$

Then if we define the functions $W_{ab}(s)$ by the equations

$$(3:15) \quad k_{ac} W_{cb} = h_{ab},$$

it follows from equations (3:14) that whenever $K \neq 0$

$$(3:16) \quad W_{ab} = K_{ac} h_{cb}.$$

Then we may state the following

LEMMA 3:4. The solutions $W_{cb}(s)$ of equations (3:15), where $h_{ab}(s)$ and $k_{ab}(s)$ are solutions of the differential equations (3:8) with initial conditions given by (3:10), are symmetric in their two subscripts:

$$(3:17) \quad W_{ab} = W_{ba},$$

and they have the derivatives

$$(3:18) \quad W_{ab}' = -K_{ac} K_{bc}.$$

To prove the first assertion sum the third set of equations (3:12) with K_{da} and then the result with K_{ab} , applying each time the identities (3:14) and (3:16). For the second part of the Lemma, differentiate equations (3:15) to obtain

$$k_{ac}' W_{cb} + k_{ac} W_{cb}' = h_{ab}'.$$

Now sum this with k_{ad} and apply the third set of equations (3:11), then (3:15), and finally the first set of equations (3:11) to obtain

$$k_{ad} k_{ac} W_{cb}' = -\delta_{bd}.$$

On summing with K_{de} and K_{ae} successively and applying equations (3:14) each time, we obtain the desired result.

4. The conjugate point. Two points P_0 and P_1 on an extremal E are said to be conjugate¹ in case there exists an orthogonal variation ξ^1 satisfying the Jacobi equations (2:23) and vanishing at the points P_0 and P_1 but not vanishing at every point of E between P_0 and P_1 . In view of the linear independence of the vectors Y_a^1 defined in the preceding section, and the method of derivation of the equations (3:8), this condition is seen to be equivalent to the existence of a solution set $h_a(s)$ of equations (3:8) vanishing at P_0 and P_1 but not identically zero between these points. But any solution $h_a(s)$ of equations (3:8) is expressible as a linear combination with constant coefficients of the fundamental sets of solutions h_{ab} and k_{ab} with initial conditions (3:10). However if $s = 0$ at P_0 so that the conditions (3:10) are imposed at P_0 , and if $h_{ab}(0) = 0$, then the coefficients of h_{ab} are all zero in the expression for $h_a(s)$ so that any such solution $h_a(s)$ is expressible in the form

$$(4:1) \quad h_a(s) = k_{ab}(s)B_b,$$

with constant coefficients B_b . From this it follows that there exists a solution $h_a(s)$ of equations (3:8) vanishing at $s = 0$ and at $s = s_1$ but not identically zero between $s = 0$ and $s = s_1$, if and only if the determinant $K(s_1) = 0$.

THEOREM 4:1. A necessary and sufficient condition for a point P_1 to be conjugate to P_0 on an extremal E is that the determinant $K(s) = |k_{ab}(s)|$ shall vanish at P_1 . The functions k_{ab} are solutions of the differential equations (3:8) with initial values (3:10) at P_0 .

Now suppose that $s_1 > 0$ is the first zero of $K(s)$, and

¹F. L. Wren, loc. cit., Section 4.

that the constants B_b are chosen to satisfy the homogeneous equations

$$(4:2) \quad k_{ab}(s_1)B_b = 0.$$

Then if $h_a(s)$ are given by equations (4:1) and $\xi^i(s)$ by equations (3:5), it follows from (3:5) and (3:10) that

$$(4:3) \quad \xi^i(0) = 0, \quad D\xi^i(0)/ds = B_a Y_a^i(0)$$

After multiplying the constants B_a by a non null constant we may assume that

$$(4:4) \quad B_a B_a = 1,$$

so that from (4:3) we have

$$(4:5) \quad g_{ij}(D\xi^i/ds)(D\xi^j/ds)|^{s=0} = 1.$$

The vectors Y_a^i were chosen arbitrarily at P_0 except for the conditions (3:3) of orthogonality and unitariness. We may therefore choose

$$(4:6) \quad Y_1^i(0) = D\xi^i(0)/ds,$$

since ξ^i has been chosen so that $\mathcal{L}_1 \xi^i = 0$ and therefore also so that $\mathcal{L}_1 D\xi^i/ds = 0$, as one sees by taking the covariant derivative. Under these circumstances it follows from equations (4:3) that

$$B_a = \delta_{1a},$$

and therefore from equations (4:2) that

$$(4:7) \quad k_{a1}(s_1) = 0.$$

If K has the rank $n - 2$ at the conjugate point P_1 , there

is but a single solution of equations (4:2) and (4:4). But if the rank of K at P_1 is less than $n - 2$, there is a solution B_a' of equations (4:2) and (4:4), independent of B_a , and for this new solution we may assume $B_1' = 0$. This determines another orthogonal solution ξ^1 of the Jacobi equations which is independent of the first and which vanishes at P_0 and P_1 . Its absolute derivative at P_0 , given by equations (4:3) with B_a' replacing B_a , lies in the space of $Y_2^1, \dots, Y_{n-1}^1$ and is therefore orthogonal to Y_1^1 . Consequently we may rotate the vectors $Y_2^1, \dots, Y_{n-1}^1$ about Y_1^1 until Y_2^1 coincides with this absolute derivative vector at P_0 , and by a procedure analogous to that leading to equations (4:7), obtain

$$(4:8) \quad k_{a2}(s_1) = 0.$$

Proceeding thus step by step we arrive at the following theorem:

THEOREM 4:2. Let the point P_1 be conjugate to P_0 on the extremal E , and let $r < n - 1$ be the rank at $s = s_1$ of the determinant K of solutions $k_{ab}(s)$ of the differential equations (3:8) satisfying the initial conditions (3:10) at P_0 . Then the vectors Y_a^1 satisfying equations (3:3) and (3:4) can always be chosen so that at P_1

$$(4:9) \quad k_{a\alpha}(s_1) = 0 \quad (\alpha = 1, \dots, n - r - 1),$$

while the remaining columns $k_{a\rho}(s_1)$ ($\rho = n - r, \dots, n - 1$) of $K(s_1)$ are linearly independent.

THEOREM 4:3. If the point P_1 is conjugate to P_0 on the extremal E , and if the vectors Y_a^1 are chosen as in the foregoing theorem, then at P_1 the functions w_{ab} defined in (3:15) become infinite for $a = b < n - r$, while for $a \geq n - r$ or $b \geq n - r$ they

are all finite. In this as in the preceding theorem r is the rank of the determinant $K(s)$ at P_1 .

In proof of this theorem we employ

LEMMA 4:1. If P_1 is conjugate to P_0 on the extremal E and the vectors Y_a^1 are chosen as in Theorem 4:2, then

A. the columns $k_{c\rho}(s_1)$ ($\rho = n - r, \dots, n - 1$) form a complete set of solutions y_c of the linear equations

$$(4:10) \quad y_c k_{c\alpha}'(s_1) = 0 \quad (\alpha = 1, \dots, n - r - 1);$$

B. if the β -th equation (4:10) is omitted, a complete set of solutions y_c of the remaining set of equations

$$(4:11) \quad y_c k_{c\delta}'(s_1) = 0 \quad (\delta = 1, \dots, \beta - 1, \beta + 1, n - r - 1)$$

is furnished by adjoining the single column $h_{c\beta}(s_1)$ to a complete set of solutions of (4:10);

C. each column $h_{c\rho}(s_1)$ is a linear combination of the columns $k_{c\rho}(s_1)$.

In the third group of the identities (3:11) set $a = \alpha$, $b = \rho$; set $a = \beta$, $b = \alpha$ in the first group; $a = \rho$, $b = \alpha$ in the first group, and evaluate at s_1 to obtain

$$(4:12) \quad k_{c\alpha}' k_{c\rho} = 0, \quad h_{c\beta} k_{c\alpha}' = \delta_{\alpha\beta}, \quad h_{c\rho} k_{c\alpha}' = 0, \quad \text{at } s_1.$$

Hence the columns $k_{c\rho}$ are solutions of (4:10). These columns are also linearly independent since the determinant K has rank r and its first $n - r - 1$ columns are all null. Equations (4:10) are also independent, since if $k_{c\alpha}' u_\alpha = 0$, then summing the second of (4:12) with u_α we find $0 = u_\beta$. This proves part A of the lemma. Part C follows then from the third set of equations (4:12).

To prove part B we observe that if, in the second set of equations (4:12), we set $\alpha = \delta \neq \beta$, then $h_{c\beta}$ is a solution of

equations (4:11). Hence it remains only to show that $h_{c\rho}$ and the columns $k_{c\rho}$ are linearly independent. But if this were not so then, because of the fact that the columns $k_{c\rho}$ are linearly independent, we should have $h_{c\rho}$ expressible as a linear combination of the $k_{c\rho}$ and hence satisfying all of equations (4:10), contrary to the second set of equations (4:12).

Now to prove the theorem, we see from the definition (3:15) or (3:16) of the functions W_{ab} that each of these is expressible as the quotient by the determinant K of a determinant, say H_{ab} , obtainable from K on replacing its a -th column by the b -th column of the matrix (h_{ab}) . Suppose first that both $a \geq n - r$ and $b \geq n - r$. Then at s_1 each of these determinants H_{ab} and K has $n - r - 1$ columns of zeros. In evaluating these indeterminate forms by the usual method, we observe that both numerator and denominator have a zero of order at least $n - r - 1$ at $s = s_1$, and after $n - r - 1$ differentiations there are, in the denominator, $(n - r - 1)!$ equal determinants with no column vanishing at s_1 . Each of these is equal to the determinant $|k_{c\alpha} k_{c\rho}|$. From part A of the lemma, application of a well-known theorem on homogeneous linear equations¹ shows the minors found from the first $n - r - 1$ columns of this determinant to be proportional, with a non-null proportionality factor, to the complementary minors. By employing a Laplace expansion therefore, we see that this determinant is equal, at s_1 , to the product of this proportionality factor by a sum of squares of minors not all of which are zero. Hence the quotient $H_{ab}/K = W_{ab}$ approaches a finite limit at s_1 when $a \geq n - r$ and $b \geq n - r$.

If $a = \alpha < n - r$ while $b = \rho \geq n - r$, then $H_{ab} = H_{\alpha\rho}$

¹H. W. Turnbull, The theory of determinants, matrices, and invariants, 1928, pp. 84-86.

has only $n - r - 2$ columns of zeros at s_1 , so that after $n - r - 2$ differentiations there are $(n - r - 2)!$ equal determinants having no column of zeros at s_1 . However the α -th column of each is the column $h_{c\rho}$ which, by part C of the lemma, is at s_1 a linear combination of the last r columns of this determinant. Hence this determinant vanishes at s_1 and an additional differentiation is required, rendering the denominator non-null as shown in the preceding paragraph. Since $W_{ab} = W_{ba}$, we have proved the last part of the theorem.

Finally if $a = b < n - r$, then $n - r - 2$ differentiations suffice to yield, in the numerator, $(n - r - 2)!$ equal determinants, each having no column of zeros at s_1 , while every determinant in the denominator has at least one such. Each determinant in the numerator has, as $n - r - 2$ of its columns, the coefficients of equation (4:11) for $a = \beta$; the β -th column is $h_{c\rho}$ and the last r columns are $k_{c\rho}$. Part B of the lemma shows, by an argument similar to that used in proving the other part of the theorem, that this determinant is non-null and the quotient H_{ab}/K therefore infinite at s_1 .

As a corollary to Theorem 4:3 we may conclude:

COROLLARY. If the vectors Y_a^1 are chosen as in Theorem 4:2, and B_α ($\alpha = 1, \dots, n - r - 1$) are any constants, not all zero, then $W_{\alpha\beta} B_\alpha B_\beta$ becomes infinite at the point P_1 conjugate to P_0 .

For if ξ^1 is an orthogonal solution of the Jacobi equations satisfying $\xi^1(0) = 0$, $D\xi^1(0)/ds = B_\alpha Y_\alpha^1$, then ξ^1 vanishes also at P_1 . Hence, according to Theorem 4:2, the Y_α^1 can be replaced by another linearly equivalent set $\bar{Y}_\alpha^1$ in the same space, with $\bar{Y}_1^1(0) = D\xi^1(0)/ds$. Then $W_{\alpha\beta} B_\alpha B_\beta = \bar{W}_{11}$ and hence becomes infinite at P_1 .

5. The focal point of a curve. We define the focal point of a curve D on an extremal E to which it is transversal, to be a point P_s such that there exists a non-null orthogonal solution of the Jacobi equations, vanishing at P_s , defining a vector tangent to D at its intersection P_0 with E , and satisfying the condition

$$(5:1) \quad [g_{1j} \xi^1 D \xi^j / ds + R^{-1} g_{1j} \xi^1 \xi^j]^{s=0} = 0,$$

where R^{-1} is the principal curvature of D at P_0 , as defined in Section 1, and s is the parameter along E having the value 0 at P_0 . Then by Theorem 2:7, the orthogonal variation ξ^1 causes the second variation I'' of the integral I along E , taken on the interval from $s = 0$ to $s = s_s$, to vanish.

After multiplying ξ^1 by a non-null constant we may suppose that

$$(5:2) \quad g_{1j} \xi^1(0) \xi^j(0) = 1.$$

If the vectors Y_a^1 are defined as in Section 3, ξ^1 can be expressed in the form (3:5) with functions $h_a(s)$ satisfying the second order differential equations (3:8). Or, employing the solutions of these equations with initial values (3:10) imposed at P_0 , we may write

$$(5:3) \quad \xi^1 = (h_{ab} A_b + k_{ab} B_b) Y_a^1,$$

where the A_b and B_b are constants. Then if we apply the initial conditions (3:10), equations (5:2) and (3:3), we find that

$$(5:4) \quad \begin{aligned} \xi^1(0) &= A_a Y_a^1(0), & D \xi^1(0) / ds &= B_a Y_a^1(0), \\ A_a A_a &= 1. \end{aligned}$$

The constants A_a are thus uniquely determined by the tangent vector ξ^1 of D at P_0 and the choice of the vectors Y_a^1 . As for

B_a , these can be found from the condition that ξ^1 shall vanish at the point P_2 for $s = s_2$. To see this we have at $s = s_2$, from equations (5:3),

$$(5:5) \quad h_{cb}(s_2)A_b + k_{cb}(s_2)B_b = 0.$$

If we sum these equations with $K_{ac}(s_2)$ and apply equations (3:16) we find

$$(5:6) \quad E_a = -W_{ab}(s_2)A_b.$$

Finally (5:1) gives, after we apply equations (5:4), (5:6), and (3:3), the following equation for the determination of the focal point:

$$(5:7) \quad W_{ab}(s_2)A_aA_b = R^{-1}.$$

Now consider the function $W_{ab}(s)A_aA_b$. This function is infinite for $s = 0$, and has a derivative

$$(5:8) \quad W_{ab}'(s)A_aA_b = -(K_{ac}A_a)(K_{bc}A_b),$$

according to Lemma 3:4. Since this is a negative sum of squares, $W_{ab}(s)A_aA_b$ is monotonically decreasing. Let us write the identity

$$W_{ab}A_aA_b = W_{\alpha\beta}A_\alpha A_\beta + 2W_{\alpha\rho}A_\alpha A_\rho + W_{\rho\sigma}A_\rho A_\sigma.$$

If the A_α are not all zero, then $W_{\alpha\beta}A_\alpha A_\beta$ becomes infinite at the point P_1 conjugate to P_0 , according to the Corollary to Theorem 4:3, but is elsewhere finite; the remaining terms in $W_{ab}A_aA_b$ are finite even at P_1 . Hence $W_{ab}A_aA_b$ becomes infinite at P_1 and assumes every finite value once and only once between P_0 and P_1 . Hence for any value R^{-1} equation (5:7) can be satisfied at one and only one point P_2 on the arc of E between P_0 and P_1 . But if every $A_\alpha = 0$, then $W_{ab}A_aA_b = W_{\rho\sigma}A_\rho A_\sigma$ remains finite and approaches

some limiting value, say $\bar{R}^{-1}$, at P_1 . Hence if $R^{-1} \geq \bar{R}^{-1}$, then equation (5:7) is again satisfied at a unique point P_2 preceding or at P_1 . Otherwise the curve D has no focal point on this arc of E . The following theorem then results from the fact that the vectors $Y_\alpha^1(0)$ define a space such that any orthogonal solution ξ^1 of the Jacobi equations vanishing at P_0 and having its absolute derivative in the space of the vectors $Y_\alpha^1(0)$ at P_0 , vanishes also at the first point conjugate to P_0 :

THEOREM 5:1. Let D be a curve transversal to an extremal E at P_0 , and let the vectors Y_a^1 ($a = 1, \dots, n - 1$) described in Section 3 be chosen so that at P_0 the vectors $Y_\alpha^1(0)$ ($\alpha = 1, \dots, n - r - 1$) define the space of the absolute derivatives of orthogonal solutions ξ^1 of the Jacobi equations which vanish at P_0 and at its first conjugate point P_1 . Then

A. if D is not orthogonal to the vectors $Y_\alpha^1(0)$, it has one and but one focal point on the arc of E between P_0 and its conjugate point P_1 , but

B. if D is orthogonal to these vectors it will have no focal point on this arc unless its principal curvature R^{-1} exceeds the finite limiting value at the conjugate point P_1 of the function $W_{\rho\sigma}(s)A_\rho A_\sigma$, where $A_\rho Y_\rho^1(0)$ defines the direction of D at P_0 , and the functions $W_{\rho\sigma}(s)$ are defined by equations (3:15).

From this theorem it is evident in any case that for an arc D with a sufficiently large curvature R^{-1} there always exists a focal point P_2 between P_0 and its conjugate point P_1 . The equation (5:7) defines the value s_2 belonging to the focal point. From the usual implicit function theorems it follows that in case A of the last theorem the single valued function $s_2(R)$ so defined increases monotonically with R and approaches the value s_1 belonging to P_1 as R^{-1} decreases from $+\infty$ to $-\infty$. In case B of the

theorem the function $s_2(R)$ increases and reaches the value s_1 for a finite value of R^{-1} . These are results analogous to those of Bliss and White in the references cited in the introduction to this paper.

6. The focal point of a variety. Let the variety V_N defined by equations (1:18) be transversal to the extremal E at P_0 , and let $(y^1, \dots, y^N) = (0, \dots, 0)$ at this point. Let the curve D on V_N defined by equations (1:19) pass through P_0 for $u = 0$. Then D is transversal to E at P_0 . Let (x, ℓ) represent an element of E . We may choose the parameter u so that at P_0

$$(6:1) \quad g_{pq} y^p y^q = 1,$$

on the arc (1:19), where g_{pq} are the coefficients of the first fundamental form for V_N as defined by equations (1:21). The vectors x_p^1 at P_0 , defined in equations (1:20), are linearly independent if, as we suppose, P_0 is not a singular point of the variety. They define the linear space S_N tangent to V_N at P_0 , and are expressible as linear combinations of the vectors $y_a^1(0)$, defined for the extremal E as in the preceding pages, of the form

$$(6:2) \quad x_p^1(0) = A_{pa} y_a^1(0).$$

The coefficients g_{pq} then have the values

$$(6:3) \quad g_{pq} |^0 = A_{pa} A_{qa}$$

at P_0 because of equations (6:2) and (3:3).

If the curve D has the direction $A_a y_a^1(0)$ at P_0 , then from equations (6:2) we may set

$$A_a y_a^1(0) = x_p^1(0) y^p(0) = A_{pa} y^p(0) y_a^1(0),$$

or

$$(6:4) \quad A_a = A_{pa} y^{p'}(0).$$

Moreover from equations (6:3) and (6:1) we find

$$(6:5) \quad A_a A_a = 1.$$

Let us define a matrix of functions $V_{pq}(s)$ by the equations

$$(6:6) \quad V_{pq}(s) = W_{ab}(s) A_{pa} A_{qb} = V_{qp}(s).$$

This matrix is symmetrical in the two indices because of the symmetry of the functions $W_{ab}(s)$. Then for the curve D

$$W_{ab}(s) A_a A_b = W_{ab}(s) A_{pa} A_{qb} y^{p'}(0) y^{q'}(0) = V_{pq}(s) y^{p'}(0) y^{q'}(0).$$

Moreover from equations (1:22) and (6:1) the curvature of the arc D has at P_0 the value

$$(6:7) \quad R^{-1} = L_{pq}(0) y^{p'}(0) y^{q'}(0).$$

Hence equation (5:7) determining the focal point of D becomes

$$(6:8) \quad [V_{pq}(s) - L_{pq}(0)] y^{p'}(0) y^{q'}(0) = 0.$$

According to Lemma 1:1, the value of $L_{pq}(0)$ does not depend upon the choice of direction of the element (x, ℓ) at any point other than P_0 , while neither $V_{pq}(s)$ nor $L_{pq}(0)$ involves derivatives of $y^p(u)$. Also equation (6:8) is homogeneous in $y^{p'}$ and hence is independent of the restriction (6:1). In view of Theorem 5:1, therefore, we may state the following theorem:

THEOREM 6:1. If V_N is a variety transversal to the extremal E at P_0 with equations of the form (1:18), and if $y^{p'}(0)$ defines a direction on V_N at P_0 , then all curves D on V_N through P_0 in the direction $y^{p'}(0)$ either have the same focal point on E , defined by the value of s satisfying equation (6:8), and lying

between P_0 and its conjugate point P_1 , or else they have none at all on this arc of E .

The set of all $y^{P'}(0)$ satisfying (6:1) is bounded and closed. There may be no direction $y^{P'}(0)$ of V_N to which there corresponds a focal point between P_0 and P_1 , in which case P_1 may itself be regarded as the focal point of the variety. But if to some $y^{P'}(0)$ there does correspond a focal point on this arc of E , then the values of s belonging to such focal points have a greatest lower bound $s_3 > 0$. A sequence of sets $y^{P'}(0)$ for which the corresponding focal values s approach s_1 will have a limiting set $y_1^{P'}(0)$ and this set will satisfy equation (6:8) with the value $s = s_3$ because of the continuity of the functions involved. By implicit function theorems the values s satisfying (6:8) for sets $y^{P'}(0)$ near to $y_1^{P'}(0)$ define a continuous and differentiable function $s(y')$. Since this function has a minimum at $y^{P'}(0) = y_1^{P'}(0)$ it follows that its derivatives with respect to $y^{P'}(0)$ all vanish at $y_1^{P'}(0)$ and hence that

$$(6:9) \quad [V_{pq}(s_3) - L_{pq}(0)]y_1^{q'} = 0.$$

These equations regarded as equations in $y^{q'}(0)$ require that their determinant vanishes:

$$(6:10) \quad |V_{pq}(s_3) - L_{pq}(0)| = 0,$$

and it follows readily that s_3 is the smallest root of this equation (6:10). The corresponding point P_3 is defined as the focal point of the variety V_N .

THEOREM 6:2. If V_N is the variety (1:18) transversal at P_0 for $y^P = 0$ to the extremal E , its focal point is the nearest of the focal points of all curves D on V_N through P_0 , if such focal points exist on E preceding the point P_1 conjugate to P_0 .

Otherwise it is to be regarded as this conjugate point P_1 itself.
In the former case the focal point P_3 corresponds to the parameter
value s_3 which is the least positive root of equations (6:10).
For this parameter value s_3 , equations (6:9) determine the direc-
tions at P_0 of all curves D through P_0 possessing as their common
focal point the point P_3 . In these equations the functions
 $V_{pq}(s)$ are linear combinations of the functions $W_{ab}(s)$ of equa-
tions (3:15) as defined by equations (6:6) and (6:2), while the
functions $L_{pq}(0)$ are defined in (1:21) at P_0 for an element
 (x, ℓ) of Finsler space whose direction is tangent to the ex-
tremal E .

The following theorem is now an immediate consequence of the preceding one and of Theorem 5:1.

THEOREM 6:3. Let the vectors $Y_\alpha^1(0)$ at a point P_0 of an
extremal arc E be determined as in Theorem 4:2. Then every
variety V_N transversal to E at P_0 and not orthogonal to all the
vectors $Y_\alpha^1(0)$ possesses a focal point P_3 on the arc of E between
 P_0 and its conjugate point P_1 . In particular, every transversal
variety of dimension $N > r$ possesses such a focal point.

